

# Can Long-Run Restrictions Identify Technology Shocks?

Christopher J. Erceg\*, Luca Guerrieri\*\*, and Christopher Gust\*\*\*

Board of Governors  
of the Federal Reserve System  
20th and C Streets, N.W., Stop 42B  
Washington, DC 20551 USA

December 2003

## Abstract

Galí's innovative approach of imposing long-run restrictions on a vector autoregression (VAR) to identify the effects of a technology shock has become widely utilized. In this paper, we investigate its reliability through Monte Carlo simulations of several relatively standard business cycle models. We find it encouraging that the impulse responses derived from applying the Galí methodology to the artificial data generally have the same sign and qualitative pattern as the true responses. However, we highlight the importance of small-sample bias in the estimated impulse responses and show that the magnitude and sign of this bias depend on the model structure. Accordingly, we caution against interpreting responses derived from this approach as “model-independent” stylized facts. Moreover, we find considerable estimation uncertainty about the quantitative impact of a technology shock on macroeconomic variables, and a corresponding level of uncertainty about the contribution of technology shocks to the business cycle.

JEL classification: E30, C52.

*Keywords:* Technology shocks, vector autoregressions, real business cycle models.

We thank James Nason, Robert Vigfusson, Chao Wei, and seminar participants at the Federal Reserve Board, Federal Reserve Bank of Atlanta, George Washington University, and the Society for Economic Dynamics 2003 Meeting. We especially thank Jon Faust for his comments. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

\* Phone: 202 452 2575; email: christopher.erceg@frb.gov

\*\* Corresponding author. Phone: 202 452 2550; email: luca.guerrieri@frb.gov

\*\*\* Phone: 202 452 2383; email: christopher.j.gust@frb.gov

# 1 Introduction

The seminal work of Blanchard and Quah (1989) has stimulated widespread interest in using vector autoregressions (VARs) that impose long-run restrictions to identify the effects of shocks. This methodology has proved appealing because it does not require a fully-articulated structural model or numerous model-specific assumptions.

One important recent application of this approach, introduced by Galí (1999), involves using long-run restrictions to identify the effects of a technology shock. The key identifying assumption in this approach is that only technology innovations can affect labor productivity in the long-run. As discussed in Galí (1999), this assumption holds in a broad class of models under relatively weak assumptions about the form of the production function. Numerous researchers have used this approach to assess how technology shocks affect macroeconomic variables, and to quantify the importance of technology shocks in accounting for output and employment fluctuations.<sup>1</sup>

While the simplicity of Galí's methodology has contributed to its broad appeal, the recent literature has suggested reasons to question whether it is likely to yield reliable inferences about the effects of technology shocks. One reason is that it is difficult to estimate precisely the long-run effects of shocks using a short data sample. Accordingly, as emphasized by Faust and Leeper (1997), structural VARs that achieve identification through long-run restrictions may perform poorly when estimated over the sample periods typically utilized. A second reason is that certain non-technology shocks, such as changes in the capital tax rate, may have permanent effects on labor productivity, thus violating Galí's key identifying assumption.<sup>2</sup>

In this paper, we critique the reliability of the Galí methodology by using Monte Carlo simulations of reasonably-calibrated dynamic general equilibrium models. In particular, we

---

<sup>1</sup>See, for example, Galí (1999), Francis and Ramey (2003), Christiano, Eichenbaum, and Vigfusson (2003), and Altig, Christiano, Eichenbaum, and Lindé (2003).

<sup>2</sup>See Altig, Christiano, Eichenbaum, and Lindé (2003). However, Francis and Ramey (2003) found that the effects of technology shocks inferred from their VAR were little affected by permanent capital tax rate changes. We also find that changes in capital tax rates play a minor role in our analysis.

compare the response of macroeconomic variables to a technology innovation derived from applying Galí’s identifying scheme with the “true” response implied by our models. We utilize two alternative models of the business cycle as the data generating process. The first is a standard real business cycle (RBC) model with endogenous capital accumulation, and includes shocks to productivity, capital and labor income tax rates, government spending, and labor supply. We consider two variants of the RBC model, one of which assumes that capacity utilization is fixed, while the other allows capacity utilization to be chosen endogenously. The second model incorporates staggered wage and price setting into the analysis. We generate Monte Carlo simulations using 45 years of quarterly data, since this sample length is similar to that used recently by most researchers.

Broadly speaking, the shocks derived from application of the Galí methodology to the simulated data “look like” true technology shocks in each of the models we consider. In particular, the point-wise mean impulse response functions (IRFs) of output, investment, consumption, and hours worked derived from the Monte Carlo simulations uniformly have the same sign and qualitative pattern as the true responses. Moreover, we find that the probability of inferring a response of output, consumption, or investment that has the qualitatively incorrect sign (even for only a few quarters) is generally low.

However, we find that the small-sample bias emphasized by Faust and Leeper (1997) poses quantitative problems for this identifying scheme. Our analysis allows us to highlight two related channels through which this bias arises. First, the slow adjustment of capital makes it hard to gauge the long-run impact of a technology shock on labor productivity, contributing to downward bias in the estimated impulse responses.<sup>3</sup> Second, the identification procedure has difficulty disentangling technology shocks from other shocks that have highly persistent, even

---

<sup>3</sup>The fact that slow adjustment of capital creates problems for the identification scheme may seem surprising given the well-known problem emphasized by Cogley and Nason (1995) that standard real business cycle models fail to generate enough endogenous persistence. However, Cogley and Nason (1995) focus on the inability of these models to generate enough positive autocorrelation in output *growth*, while our emphasis is on the *level* of labor productivity.

if not permanent, effects on labor productivity (such as labor supply or tax rate shocks).<sup>4</sup> As a result, even in the absence of shocks that would violate Galí’s long-run identifying assumption, the estimated technology shock may incorporate a sizeable non-technology component. Accordingly, the bias in the estimated response of a given variable to a technology shock depends on the relative magnitude of technology and non-technology shocks, and on its response to non-technology shocks.

The slow adjustment of capital mainly accounts for the sizeable downward bias in the mean response of output derived from the Monte Carlo simulations in each of the models considered. Moreover, given substantial spread in the distribution of the impulse responses, we find that the probability that a researcher would estimate an output impulse response that lies uniformly more than 33 percent away from the true response (for the first four quarters following the shock) exceeds 40 percent in two of the three model variants. This bias and spread in estimating the impulse responses also implies that the structural VAR may produce poor estimates of the contribution of technology shocks to business cycle fluctuations. In the model with nominal rigidities, we find there is almost a 50 percent chance of underpredicting the volatility of HP-filtered output due to technology shocks by 50 percent or more.

The inability of the structural VAR to disentangle technology shocks from non-technology shocks accounts for the surprising result that the bias in the response of hours and investment varies substantially across the models considered. For example, the mean response of investment overstates the true response by nearly 70 percent (on average over the first 12 quarters) in the RBC model with variable capacity utilization, reflecting that expansionary labor supply and tax rate shocks generate a sharp initial rise in investment spending. By contrast, the mean response of investment understates the true response by 45 percent in the model with nominal rigidities, reflecting that the monetary rule induces investment to contract initially in response to the same labor supply/tax shocks. Thus, the bias may differ substantially across models, because the way in which a variable responds to non-technology shocks is model-dependent.

---

<sup>4</sup>In this respect, our paper is similar to an earlier literature emphasizing that the measured Solow residual is contaminated by aggregate demand disturbances. See, for example, Evans (1992) and references therein.

Overall, Galí’s methodology appears to offer a fruitful approach to uncovering the effects of technology shocks, but our analysis suggests important caveats in interpreting the results. While the key long-run identifying assumption employed by Galí may hold across a broad class of models, the ability of the SVAR to identify technology shocks may be sensitive to specific features of the data-generating process. Accordingly, our analysis suggests that empirical estimates of the effects of technology shocks should not be regarded as model-independent stylized facts. In particular, a researcher should not simply compare the impulse responses to a technology shock from a theoretical model to those obtained from applying Galí’s identifying scheme to actual data. Rather, it is important to adjust for small-sample bias, the size and direction of which depends on the model under consideration.

This paper is organized as follows. Section 2 outlines our baseline RBC model and describes the calibration. Section 3 reviews the Galí identification scheme. Section 4 reports our results for both versions of the RBC model. Section 5 discusses the results for the model with nominal rigidities, and Section 6 examines the coverage of conventional bootstrap confidence intervals. Section 7 concludes.

## 2 The RBC Model

We begin by outlining a relatively standard real business cycle model that includes variable capacity utilization as in Burnside and Eichenbaum (1996). The “textbook” RBC model with fixed capacity utilization emerges as a special case.

### 2.1 Household Behavior

The utility function of the representative household is

$$E_t \sum_{j=0}^{\infty} \beta^j \left\{ \log(C_{t+j}) + \chi_{0t+j} \frac{(1 - N_{t+j})^{1-\chi}}{1-\chi} \right\}, \quad (1)$$

where the discount factor  $\beta$  satisfies  $0 < \beta < 1$  and  $E_t$  is the expectation operator conditional on information available at time  $t$ . The period utility function depends on consumption,  $C_t$ ,

leisure,  $1 - N_t$ , and a stochastic shock,  $\chi_{0t}$ , that may be regarded as a shock to labor supply. We assume that this labor supply shock evolves according to:

$$\log(\chi_{0t}) = (1 - \rho_\chi) \log(\chi_0) + \rho_\chi \log(\chi_{0t-1}) + \sigma_\chi \epsilon_{\chi t}, \quad (2)$$

where  $\chi_0$  denotes the steady state value of  $\chi_{0t}$  and  $\epsilon_{\chi t} \sim N(0, 1)$ .

The representative household's budget constraint in period  $t$  states that its expenditure on consumption and investment goods ( $I_t$ ) and net purchases of bonds  $B_{t+1}$  must equal its after-tax disposable income:

$$C_t + I_t + \frac{1}{1+r_t} B_{t+1} - B_t = (1 - \tau_{Nt}) W_t N_t + \Gamma_t + T_t + (1 - \tau_{Kt}) R_{Kt} v_t K_t + \tau_{Kt} \delta K_t - 0.5 \phi_K K_t \left( \frac{I_t}{K_t} - \widehat{\delta} \right)^2. \quad (3)$$

The household earns after-tax labor income of  $(1 - \tau_{Nt}) W_t N_t$ , where  $\tau_{Nt}$  is a stochastic tax on labor income, and also receives an aliquot share of firm profits  $\Gamma_t$  and a lump-sum government transfer of  $T_t$ . The household leases capital services to firms at an after-tax rental rate of  $(1 - \tau_{Kt}) R_{Kt}$ , where  $\tau_{Kt}$  is a stochastic tax on capital income. Capital services depend multiplicatively on the size of the household's capital stock  $K_t$ , and on the household's choice of a utilization rate  $v_t$ . The household receives a depreciation writeoff of  $\tau_{Kt} \delta$  per unit of capital (where  $\delta$  is the steady state depreciation rate of capital). The household also faces adjustment costs for changing the ratio of its investment to its capital stock relative to the steady-state level of  $\widehat{\delta} = \delta e^{\mu_z}$ , where  $\mu_z$  denotes the steady-state logarithmic growth rate of technology.

Purchases of investment goods augment the household's capital stock according to the transition law:

$$K_{t+1} = (1 - \delta(v_t)) K_t + I_t. \quad (4)$$

The depreciation rate of capital is assumed to depend positively on the household's choice of a utilization rate<sup>5</sup>, and we assume that this function is of the isoelastic form:

$$\delta(v_t) = \delta v_t^\phi, \quad (5)$$

---

<sup>5</sup>See Greenwood, Hercowitz, and Huffman (1988) or Burnside and Eichenbaum (1996).

where  $\phi > 1$ . In every period  $t$ , the household maximizes utility (1) with respect to its consumption, labor supply, investment, (end-of-period) capital stock, capital utilization rate, and real bond holdings, subject to its budget constraint (3), and the transition equation for capital (4).

## 2.2 Firms

The representative firm uses capital services ( $v_t K_t$ ) and labor to produce a final output good that can either be consumed or invested. This firm has a constant returns-to-scale Cobb-Douglas production function of the form:

$$Y_t = (v_t K_t)^\theta (Z_t N_t)^{1-\theta}, \quad (6)$$

Here,  $Z_t$  is the exogenous process for technology whose law of motion is governed by:

$$\log(Z_t) - \log(Z_{t-1}) = \mu_z + \sigma_z \epsilon_{zt}, \quad (7)$$

where  $\epsilon_{zt} \sim N(0, 1)$ .

The firm purchases capital services and labor in perfectly competitive factor markets, so that it takes as given the rental price of capital services  $R_{Kt}$  and the aggregate wage  $W_t$ .<sup>6</sup> Since the firm behaves as a price taker in the output market as well as in factor markets, the following efficiency conditions hold for the choice of capital services and labor:

$$\frac{MC_t}{P_t} = \frac{W_t}{MPL_t} = \frac{R_{Kt}}{MPK_t} = 1. \quad (8)$$

## 2.3 Government

Some of the final output good is purchased by the government, so that the market-clearing condition is:

$$Y_t = C_t + I_t + G_t. \quad (9)$$

---

<sup>6</sup>Note that in our model, households determine the aggregate of capital services by their choice of the utilization rate of capital and level of the capital stock. Correspondingly, firms are indifferent as to whether households produce more capital services by increasing the stock of capital, or by increasing the utilization rate of a fixed stock of capital.

Government purchases are assumed to have no direct effect on the utility function of the representative household. We also assume that government purchases are exogenous and share a common trend with the level of technology. Therefore, we define  $g_t = G_t/Z_t$ , which evolves according to:

$$\log(g_t) = (1 - \rho_g) \log(g) + \rho_g \log(g_{t-1}) + \sigma_g \epsilon_{gt}, \quad (10)$$

where  $g$  denotes the steady state value of  $g_t$  and  $\epsilon_{gt} \sim N(0, 1)$ .

The government's budget is balanced every period, so that total taxes – which include both distortionary taxes on labor and capital income – equal the sum of government purchases of the final output good and net lump-sum transfers to households.<sup>7</sup> Hence, the government's budget constraint at date  $t$  is:

$$T_t + G_t = \tau_{Nt} W_t N_t + \tau_{Kt} (R_{Kt} - \delta) K_t. \quad (11)$$

The tax rates on capital and labor are assumed to be exogenous and evolve according to:

$$\tau_{it} = (1 - \rho_{\tau_i}) \tau_i + \rho_{\tau_i} \tau_{it-1} + \sigma_{\tau_i} \epsilon_{\tau_i t}, \quad (12)$$

where  $\tau_i$  is the steady state tax rate and  $\epsilon_{\tau_i t} \sim N(0, 1)$  for  $i = K, N$ .

## 2.4 Solution and Calibration

To analyze the behavior of the model, we first apply a stationary-inducing transformation to those real variables that share a common trend with the level of technology. This entails detrending real GDP, the GDP expenditure components, and the real wage by  $Z_t$  and the capital stock,  $K_t$ , by  $Z_{t-1}$ . We then compute the solution of the model using the numerical algorithm of Anderson and Moore (1985), which provides an efficient implementation of the solution method proposed by Blanchard and Kahn (1980).

We calibrate two versions of the RBC model: in one case, capacity utilization is simply fixed (the FCU version) so that  $\delta(v_t) = \delta \forall t$ , while in the alternative case it is allowed to

---

<sup>7</sup>The assumption of a balanced budget is not restrictive given the availability of lump-sum taxes or transfers.



vary (the VCU version). Both versions are calibrated at a quarterly frequency. We begin by discussing the FCU calibration. As reported in Table 1, we assume  $\beta = 1.03^{-0.25}$ , consistent with a steady-state annualized real interest rate of 3 percent. The utility function parameter  $\chi_0$  is set so that steady-state employment (N) comprises one-third of the household's time endowment. The parameter  $\chi$  is set to 3, which implies a Frisch elasticity of labor supply of  $2/3$ . Such an elasticity is lower than if preferences were logarithmic in leisure, but within the range of most empirical estimates.<sup>8</sup> The capital share parameter  $\theta$  is set to 0.35, while  $\delta = 0.02$ , consistent with an annual depreciation rate of 8 percent. Because the model without variable capacity utilization tends to underpredict investment volatility even when there are no adjustment costs, we set  $\phi_K = 0$  in this case.

For the parameters governing the two tax rate series, we estimated equation (12) using OLS after constructing these tax rates series based on U.S. data from 1958-2002 following the methodology described in Jones (2002).<sup>9</sup> Our estimates implied  $\tau_K = 0.38$ ,  $\rho_{\tau_K} = 0.97$ , and  $\sigma_{\tau_K} = 0.008$  for the capital tax rate, and  $\tau_N = 0.22$ ,  $\rho_{\tau_N} = 0.98$ , and  $\sigma_{\tau_N} = 0.0052$  for the labor tax rate.

We constructed a quarterly measure of the capital stock by using data on investment, an initial value for the capital stock, and the capital accumulation equation (4).<sup>10</sup> Defining the Solow residual as:

$$S_t = \frac{Y_t}{K_t^\theta N_t^{1-\theta}}, \quad (13)$$

and noting that  $Z_t = S_t^{\frac{1}{1-\theta}}$ , we obtained a time series for  $Z_t$ . Accordingly, we equated  $\mu_z$  with the sample mean of the logarithmic growth rate of technology, and  $\sigma_z$  with the sample standard deviation. This yields  $\mu_z = 0.0037$  and  $\sigma_z = 0.0148$ .

Using government spending data and our derived  $Z_t$  series, we estimated a first order autoregression for the scaled government spending shock  $g_t$  (allowing for a linear time trend),

---

<sup>8</sup>See, for example, Pencavel (1986), Killingsworth and Heckman (1986), and Pencavel (2002).

<sup>9</sup>Following Appendix B in Jones (2002), we used quarterly data collected by the Bureau of Economic Analysis.

<sup>10</sup>We used quarterly NIPA data from 1958-2002 on real gross private domestic investment and assumed an initial value of capital three times larger than annualized output in the nonfarm business sector.

and found  $\rho_g = 0.95$  and  $\sigma_g = 0.016$ .<sup>11</sup> We set  $g$  so that the ratio of government spending to output is 20%, a value consistent with the share of government consumption to GDP for the United States based on NIPA data.

In the absence of labor-supply shocks, our calibrated RBC model would significantly underestimate the volatility in hours worked – a familiar problem in the real business cycle literature. To see this, Table 2 compares the second moments of several key variables that are implied by our model with their sample counterparts based on U.S. data. As shown in the column labelled “ $\sigma_\chi = 0$ ”, the model significantly understates the ratio of the standard deviation of HP-filtered hours to the standard deviation of HP-filtered output. For our benchmark calibration, we address this issue by incorporating labor supply shocks.<sup>12</sup> However, we also consider the alternative approach following Hansen (1985) and Rogerson (1988) of specifying that utility is linear in leisure (i.e.,  $\chi = 0$ ). In this case, which we call “High LSE” calibration, we do not incorporate labor supply shocks.

In our benchmark calibration, we add a labor-supply shock  $\chi_{0t}$  with an innovation variance  $\sigma_\chi$  that is estimated by the method of moments. In particular, we estimate  $\sigma_\chi$  so that it allows the model to exactly match the observed ratios of the standard deviation of hours worked relative to output, where both the hours and output data have been HP-filtered. We set the persistence parameter of the labor-supply shock  $\rho_\chi = 0.98$ , identical to that of the labor-tax rate. Given that the labor-supply shock operates exactly like a labor-tax rate shock, the inclusion of the former is tantamount to scaling up the magnitude of the labor tax rate shock. One reason we chose such a high value for  $\rho_\chi$  is that it helps the model generate more persistence in output growth.<sup>13</sup> Moreover, it seems plausible that empirically relevant labor supply shocks, including those driven by demographic change and participation decisions have persistent effects.

---

<sup>11</sup>For this regression, we used quarterly NIPA data on real government consumption and investment expenditures from 1958-2002. Our estimates for  $\rho_g$  and  $\sigma_g$  are close to Burnside and Eichenbaum (1996).

<sup>12</sup>Others who have followed this approach include Hall (1997), Shapiro and Watson (1988), and Parkin (1988).

<sup>13</sup>Even with  $\rho_\chi = 0.98$ , the model was unable to match the persistence of output growth in U.S. data. For a discussion of this well-known problem with RBC models, see Cogley and Nason (1995).

Table 2 shows the selected moments for the benchmark FCU and the High LSE calibrations. A comparison of the model’s implications for the volatility of output, investment, and consumption to the corresponding sample moments suggests that these calibrated models perform quite well on these dimensions, even though they were not calibrated specifically to match these moments.

Next, we consider the calibration of the VCU version of the model. Upon normalizing  $\phi$  so that capacity utilization equals unity ( $v = 1$ ) in the non-stochastic steady state, the first-order condition for choosing capacity may be expressed as  $(1 - \tau_K)R_K = \phi\delta$ . This allows us to solve for  $\phi$  as a function of other model parameters as follows:

$$\phi = \frac{\mu_z/\beta - 1}{\delta} + (1 - \tau_K) = 1.18. \quad (14)$$

In the VCU version, variation in the Solow residual reflects both changes in technology and movements in the unobserved level of capacity utilization in response to all of the underlying shocks. We used the method of moments to infer the technology and labor supply innovations that allow the model to match exactly the empirical volatility of the Solow residual growth rate and the HP-filtered ratio of hours worked relative to output (the same moments as in the benchmark FCU calibration). In addition, we chose  $\phi_K$  so that the model matches the observed ratio of the standard deviation of HP-filtered investment to the standard deviation of HP-filtered output. In the column labelled “VCU Benchmark”, Table 2 shows the values of  $\sigma_z$ ,  $\sigma_\chi$ ,  $\phi_K$  from this calibration exercise. With the exception of these parameters, our procedure assumes that the other calibrated parameters are the same as in the benchmark FCU calibration.<sup>14</sup>

### 3 The SVAR Specification

In this section, we outline the estimation procedure that a researcher is presumed to follow given a single realization of data. The structural VAR (SVAR) a researcher would estimate

---

<sup>14</sup>We recomputed the Solow residual to take account of time-varying depreciation on the estimated capital stock using a recursive procedure similar to that of Burnside and Eichenbaum (1996). However, we found that the volatility of the Solow residual was the same as in the case in which depreciation was fixed.

takes the form:

$$A(L)X_t = A_0e_t, \tag{15}$$

where  $A(L) = I - A_1L - \dots - A_pL^p$ , and  $A_i$  for  $i = 1, 2, \dots, p$  is a square matrix of reduced-form parameters;  $L$  is the lag operator, and  $X_t$  and  $e_t$  are vectors of endogenous variables and disturbances, respectively.  $X_t$  contains the log difference of average labor productivity, the log of hours worked, the log of consumption-to-output ratio, and the log of investment-to-output ratio. All variables are expressed as a deviation from the model's nonstochastic steady state, and average labor productivity is defined as  $Y_t/N_t$ .<sup>15</sup> The lag length,  $p$ , is chosen by using the information criterion in Schwarz (1978), where  $p \in \{1, 2, \dots, 10\}$ .

The inclusion of average labor productivity growth in  $X_t$  is standard in the empirical literature using VARs to identify technology shocks. While the empirical literature is divided on whether hours worked are best included in levels or differences, the former specification is selected, because the DGE model implies that hours are stationary in levels. The ratios of investment and consumption to output are included in the VAR, because Christiano, Eichenbaum, and Vigfusson (2003) have found these variables to be important in controlling for omitted-variable bias when using U.S. data. Including these variables also allows us to investigate the ability of the VAR to adequately capture their dynamic responses.

The identification of the technology shock is achieved in the following way. First, it is assumed that the innovations are orthogonal and have been normalized to unity so that

$$Ee_te_t' = I. \tag{16}$$

Denote the first element of  $e_t$  as  $e_{zt}$ , the technology shock identified by the VAR. Following Galí (1999), a researcher would then impose that the technology shock is the only shock that

---

<sup>15</sup>In the appendix, we show that, for the benchmark FCU and VCU calibrations, the model's log-linear decision rules imply that the variables in  $X_t$  can be represented as a VARMA(4,5). As discussed in the appendix, we verify numerically that this VARMA(4,5) process is invertible and is a fundamental representation for  $X_t$ . Consequently, the VAR that we study is not subject to the criticism of Lippi and Reichlin (1993).

can affect the level of productivity in the long run, an assumption that is consistent with the models we consider. Thus, letting  $R(L) = A(L)^{-1}$ , it follows that

$$[R(1)A_0]_{1j} = 0 \quad \text{for } j = \{2, 3, 4\}. \quad (17)$$

Here,  $R(L)$  is the reduced-form moving average representation of the VAR given by

$$R(L) = \sum_{i=0}^{\infty} R_i L^i, \quad (18)$$

where  $R_i$  is a  $4 \times 4$  matrix of parameter estimates and  $R_0 = I$ . The restrictions associated with equation (17) are imposed through a Cholesky decomposition after estimating  $A(L)$  using least squares. This decomposition is used to solve for the first row of  $A_0$  given that  $R(1) = A(1)^{-1}$ . No attempt is made to identify the non-technology shocks.

In our Monte Carlo study, we generate 5000 data samples from the relevant DGE model, and apply the estimation strategy discussed above to each sample. Every data sample consists of 180 quarterly observations.

## 4 Estimation results

In this section, we first discuss estimation results for the version of the RBC model in which capacity utilization is fixed, and then consider the version with variable capacity utilization.

### 4.1 Results for the FCU Calibration

Figure 1 reports the response of labor productivity, hours worked, consumption, investment, and output to a technology shock for the benchmark FCU calibration. In each panel, the solid lines show the true responses from the DGE model. The innovation occurs at date 1 and has been scaled so that the level of labor productivity rises by one percent in the long run.

The dashed lines show the point-wise mean of the impulse responses derived from applying the SVAR estimation strategy to the 5000 artificial data samples (the point-wise median

response is nearly identical).<sup>16</sup> The dotted lines show the 90 percent point-wise confidence interval of the SVAR's impulse responses.

As shown in Figure 1, the mean responses of labor productivity, consumption, investment, and output have the same sign and qualitative pattern as the true responses. As indicated by the point-wise confidence intervals, the SVAR is likely to give the appropriate sign of the response for these variables. For hours worked, the mean estimate is also qualitatively in line with the true response; however, the confidence interval is very wide, indicating that negative estimates are almost as likely as positive ones.

Quantitatively, the SVAR does not perform as well. As seen in Figure 1, the mean responses of the SVAR systematically underestimate labor productivity, consumption, investment, and output, while overestimating hours worked. To gauge the size of the bias, the top row of Table 3 reports the average absolute percent difference between the mean response and the true response over the first twelve quarters for each of the variables.<sup>17</sup> As reported in the first row of Table 3, labor productivity is underestimated by the SVAR by 32% on average over the first 12 quarters after the innovation to technology, while output is underestimated by 22%.<sup>18</sup> We defer our explanation of these results to the next subsection.

While useful for illustrating the bias associated with the SVAR's estimates, the relative distance measure does not capture the uncertainty that a researcher confined to a single draw of the data would confront. After all, the impulse response derived using a single realization of the data may diverge substantially from the mean. Accordingly, we consider an alternative measure of how well the SVAR's point estimates of the impulse responses match the truth. For

---

<sup>16</sup>We scale up the technology innovation derived from the SVAR by the same constant factor as applied to the true innovation.

<sup>17</sup>For variable  $i$ , this measure is defined as  $rd_i^m = \frac{1}{12} \sum_{l=1}^{12} |rd_{l,i}^m|$  where  $rd_{l,i}^m = \frac{\hat{d}_{l,i}^m - d_{l,i}^*}{d_{l,i}^*}$ , and  $d_{l,i}^*$  and  $\hat{d}_{l,i}^m$  denote the DGE model's impulse response and the SVAR's point-wise mean response to a technology shock for the  $i^{th}$  variable in  $X_t$  at lag  $l$ , respectively.

<sup>18</sup>Given the model structure, it is appropriate to include hours in log-levels in the VAR. However, we also examined a specification with hours in log-difference and found that small-sample biases remained substantial. These results are shown in Table 3 for the FCU calibration.

variable  $i$ , this measure is defined as

$$\hat{P}_i(\frac{1}{3}) = P(|rd_{l,i}| \geq \frac{1}{3}) \quad \forall l \in \{1, 2, \dots, N\}, \quad (19)$$

where  $rd_{l,i} = \frac{\hat{d}_{l,i} - d_{l,i}^*}{d_{l,i}^*}$  and  $\hat{d}_{l,i}$  denotes the estimated impulse response for the  $i^{th}$  variable at lag  $l$  for a given draw of data, and  $d_{l,i}^*$  denotes the response from the DGE model. In words,  $\hat{P}_i(\frac{1}{3})$  is the probability that the SVAR produces an impulse response that lies at least 33 percent above or below the true response for all lags between 1 and  $N$ , which we call a “large” error. Tables 4, 5, and 6 show these probabilities for  $N$  equal to four, eight, and twelve, respectively (as noted below, we define the measure of a large error for hours worked differently). As shown in the top row of Table 4, the probability of a large error over the first year is 34% for labor productivity and 22% for output. Furthermore, we found that nearly all of the large misses of the SVAR’s impulse responses for output and labor productivity were the result of underpredicting the true response. Given the very strict criterion that only counts impulse response functions that lie uniformly outside the 33 percent band, our results suggest considerable estimation uncertainty about the quantitative effects of a technology shock.<sup>19,20</sup>

While the probability of underestimating labor productivity, consumption, output, and investment is substantial, the probability of inferring a qualitatively incorrect sign for several quarters is very low (not reported). However, it is interesting to assess the probability of inferring a response of hours worked that is qualitatively incorrect over an extended period, particularly in light of the significant attention recent research has devoted to this question.<sup>21</sup>

---

<sup>19</sup>We also considered

$$\bar{P}(\frac{1}{3}) = \frac{1}{N} \sum_{l=1}^N P(|\hat{r}d_{l,i}| \geq \frac{1}{3}). \quad (20)$$

This measure is the average probability that the SVAR produces an error at least a third as large as the true response for variable  $i$ . Given the weaker nature of this condition, relative to (19), the probability of a large error for a given  $N$  was considerably higher.

<sup>20</sup>In the appendix, we show that our results are not substantially influenced by using a fixed lag length instead of the Schwarz criterion.

<sup>21</sup>Galí (1999) found that hours worked declined after a technology shock, and argued that this was problematic for the real business cycle paradigm, in which an expansion of hours plays a key role. But this result has been

Accordingly, for hours worked, Tables 4, 5, and 6 report the probability that the estimated response of hours worked is uniformly negative over the first 4, 8, and 12 quarters, respectively. As shown in Figure 1, the true response of hours is positive, and there is upward bias in the mean estimated response. Nevertheless, there is a 23% chance a researcher would conclude that hours worked fell in the first year after a technology shock.

Another interesting question is whether the SVAR produces reliable estimates of the contribution of technology shocks to output fluctuations over the business cycle. Table 7 shows the relative contribution of technology shocks to HP-filtered output defined as  $RC_z = \sigma_{y|z}^2 / \sigma_y^2$  where  $\sigma_y^2$  denotes the unconditional variance of HP-filtered output in the model and  $\sigma_{y|z}^2$  is the variance of HP-filtered output conditional on only technology shocks. As shown in the first row of Table 7, technology shocks account for roughly three quarters of the variance of output at business cycle frequencies in the benchmark FCU version of the model. However, the SVAR's median estimated contribution of technology shocks is only about two-thirds as large. Moreover, there is a 30 percent chance that the SVAR underreports the importance of technology shocks by 50 percent or more.<sup>22</sup>

## 4.2 Interpreting the Bias

Here we begin by providing a statistical interpretation of the bias in the point-wise mean of the impulse responses that indicates it is largely attributable to limited sample size. We then provide an economic interpretation.

The bias in the mean response of a variable can be decomposed into two sources. The first source reflects that the finite-ordered VAR used in estimation is an imperfect approximation to the infinite-ordered VAR representation implied by the true model.<sup>23</sup> This source of bias

---

the subject of considerable contention. Francis and Ramey (2003) corroborate Galí's result, while Christiano, Eichenbaum, and Vigfusson (2003) conclude that hours worked rise following a technology innovation.

<sup>22</sup>Most research has found that technology shocks play a small role in driving output fluctuations over the business cycle. A notable exception is Fisher (2002), who attempts to discriminate between multi-factor productivity and investment-specific technology shocks.

<sup>23</sup>In the appendix, we show that the linear dynamics of the DGE model can be expressed as an invertible



persists asymptotically and was discussed by Cooley and Dwyer (1998) in a similar context.<sup>24</sup> The second source of bias is attributable to limited sample size.

We illustrate this decomposition for the response of labor productivity in the lower right panel of Figure 1. The mean bias in labor productivity is represented by the solid line, labelled “total error”. The component of the bias that persists asymptotically is represented by the dotted line, labelled “ $\zeta$  error”, while the residual between the total error and  $\zeta$  error is the small sample bias. As seen in the figure, the asymptotic bias only comprises a tiny fraction of the overall bias in the mean response. Similarly, we found that the asymptotic bias is a small component of the total bias in the responses of the other variables considered in our analysis (not shown).

The magnitude of the small sample bias is largely attributable to the difficulty in precisely estimating the long-run response of variables to the innovations in the VAR. To understand this, we followed Faust and Leeper (1997) by decomposing the small-sample bias into two parts. Noticing that equation (15) can be expressed as:

$$X_t = A(L)^{-1}A_0e_t = R(L)A_0e_t, \quad (21)$$

it is evident that the response of  $X_t$  to the underlying innovations,  $e_t$ , is influenced both by the reduced-form moving average terms and by the identifying restrictions as reflected in  $A_0$ . Therefore, we can think of one part of the bias as reflecting the small-sample error in estimating the reduced-form moving average terms,  $R(L)$ . The second part emphasizes the error associated with estimating the long-run response of  $X_t$  to shocks, which translates into error in the estimation of  $A_0$ . This occurs because  $A_0$  is implicitly a function of the long-run responses as can be seen from equation (17). A more detailed description of the decomposition is given in the appendix.

Returning to the lower right panel of Figure 1, the dashed line, labelled “R error”, depicts the first type of small sample error, while the dashed-dot line labelled “ $\alpha$  error” shows the second type of error. It is clear that most of the small-sample bias in labor productivity is initially

---

VARMA(4,5). This process also has an infinite-ordered VAR representation.

<sup>24</sup>Cooley and Dwyer (1998) also emphasize other aspects of misspecification in structural VARs.

attributable to error in estimating the long-run restriction. Eventually, however, imprecision in estimating the long-run responses has a roughly commensurate effect on each component.

We now proceed to give an economic interpretation of the bias in the impulse responses. This bias reflects two related factors. First, the slow adjustment of capital makes it hard to estimate the long-run impact of a technology shock on labor productivity, contributing to downward bias in the estimated impulse responses. Second, the SVAR has difficulty disentangling technology shocks from highly persistent non-technology shocks, so that the estimated technology shock may incorporate a sizable non-technology component. The second source of bias has more pronounced effects on the estimated responses to a technology shock as the relative magnitude of non-technology shocks rises, and as the non-technology shocks become more persistent.

To demonstrate the role that the slow adjustment of capital plays in our analysis, Table 3 shows the effect of increasing the capital adjustment cost parameter by setting  $\phi_K = 100$  (leaving other parameters unchanged). This modification induces capital and labor productivity to converge much more slowly to their new long-run level following a technology shock. The downward bias in the mean response of labor productivity, output, consumption, and investment is markedly accentuated in this case. Moreover, as shown in Tables 4 to 6, the increase in the bias translates into a greater probability of making a large miss in estimating the true response of each of these variables.

We illustrate the second source of bias through three experiments. In our first experiment, we reduce the innovation variance of the technology shock to one-third of its value in the benchmark FCU calibration, so  $\sigma_z = 0.0049$  (again, leaving other parameters unchanged). As shown in Figure 2, the downward-bias in the response of labor productivity is larger in this case than in the benchmark FCU calibration. By contrast, the upward bias in hours worked is more pronounced, while the mean response of investment now exhibits substantial upward bias (see also Table 3). These changes across calibrations reflect that the estimated technology innovation inherits a larger non-technology shock component as the relative variance of the technology shock declines. Figure 3 plots the true impulse responses to a (stimulative)

labor supply innovation,  $\epsilon_\chi$ , a (negative) innovation in the capital-tax rate,  $\epsilon_{\tau_K}$ , and to the technology innovation used in this calibration ( $\sigma_z = 0.0049$ ). The non-technology shocks have a persistent effect on labor productivity, but the size of their effect is much smaller than that of a technology shock. Relative to their effect on labor productivity, the non-technology shocks exert a proportionately larger influence on hours worked and investment than a true technology shock. Accordingly, the estimated responses to a technology innovation shown in Figure 2 look more like responses to a labor supply shock (or labor-tax rate shock), which is the dominant non-technology shock.

A second way to demonstrate the important influence that the non-technology shocks have on the SVAR’s estimated technology shock is to lower their persistence. In the rows of Tables 3 to 6 labelled “with lower persistence”, we show the effects of halving all of the AR(1) parameters that govern the persistence of the non-technology shocks from their benchmark values. As seen in Table 3, the (percentage) distance between the mean and the true response narrows for all variables. The frequency of large misses also declines, though it is important to caution that the decline in the frequency of large misses reported in the tables is only partly attributable to improved estimation performance. In particular, because the estimated responses tend to be more “bouncy” as the persistence of the non-technology shocks declines, it becomes less likely that responses will lie *uniformly* outside of the 33 percent band.

Finally, we analyze the alternative calibration of the FCU model that assumes that the Frisch elasticity of labor supply is infinite ( $\chi = 0$ ). The solid lines in Figure 4 plot the response of labor productivity, hours worked, consumption, investment, and output to a technology shock in this case. There is a slight downward bias in the response of hours worked, rather than the pronounced upward bias in the benchmark FCU calibration. This disparity mainly reflects that labor supply shocks are not included in this alternative case, so that the estimated technology shock incorporates a smaller non-technology shock component.

### 4.3 Additional Sensitivity Analysis

Our consideration of alternative parameterizations of the FCU model demonstrates one of our key results: namely, even if the long-run identifying assumption used by Galí holds exactly in our model, the estimated responses to a technology innovation will be biased in small samples due to persistent non-technology shocks. Moreover, as highlighted in our investigation of alternative models below, because the relative magnitude and the effects of non-technology shocks depend on the particular model structure, the bias in the estimated response to a technology shock is also model-dependent. Before proceeding to this investigation of alternative models, it is interesting to consider two more experiments. First, we assess the importance of changes in capital tax rates. Second, we examine the role of sample length in influencing our results.

Permanent changes in capital tax rates have been recognized as a potential problem for the Galí identification scheme, as they would have permanent effects on labor productivity (thus violating the long-run identifying assumption). However, in the models we consider, the volatility of changes in the capital tax rate is too small for this to be an empirically relevant issue. As shown in the rows of Tables 3-6 labelled “with  $\rho_{\tau_K} \approx 1$ ”, we find little difference in results when the persistence of the capital tax rate shock is set arbitrarily close to unity.<sup>25</sup> Moreover, our results would also change very little if we essentially “zeroed-out” the capital tax rate shock by setting its innovation variance equal to 1/10 of its baseline value. The results from this latter case are shown under the label “with  $\sigma_{\tau_K} \approx 0$ ” in the tables.

Table 8 documents the performance of the SVAR using data samples of different lengths, generated under the benchmark FCU calibration. In practice, researchers might be limited to samples shorter than 180 quarterly observations, or might choose to discard part of the sample available because of structural breaks. In the row labelled “120”, which corresponds to 30 years of quarterly data, we report the probabilities of large misses over the first four quarters following the shock. Not surprisingly, our results suggest that the problems documented above are compounded by reducing the length of the estimation sample. Interestingly, the table

---

<sup>25</sup>Our solution procedure linearizes around a unique steady state, and thus presumes all non-technology shocks are stationary. Accordingly, we set  $\rho_{\tau_K} = 0.9999$ .

shows that there would be a sizeable chance of making large errors even with 100 years of data. For instance, the probability that the response of labor productivity would be estimated uniformly outside a 33% band around the true response remains as high as 19%. Only when the estimation sample includes 1000 quarterly observations do most of the probabilities of large misses drop below 10%.

## 4.4 Results for the VCU Calibration

Figure 5 shows the effects of a technology shock for the VCU calibration, which allows utilization to vary. While the mean estimated response of labor productivity is still below the true response, there is now a large overestimation of the responses of hours worked and investment. These large misses are confirmed in Table 3, whose row labelled “Benchmark VCU” shows that the mean estimated response of investment is nearly 70% above the DGE model’s response on average over the first twelve periods, while the mean estimated response of hours is over five times the true response.

This large overestimation of the response of hours worked and investment derives from the lower variance of the technology shock in the VCU calibration. As discussed earlier, this causes the estimated technology shock to embody a larger non-technological component. As a result, the estimated responses of hours worked and investment to a technology shock begin to resemble their responses to a tax cut or positive labor-supply shock.<sup>26</sup>

Table 4 shows that the chance of a researcher falsely concluding that hours worked falls after a technology shock is still a non-negligible 13%, despite the upward bias of the response of hours worked shown in Figure 5. However, in this case, it is more likely that a researcher actually overestimates the response of hours. Accordingly, we computed the probability that the

---

<sup>26</sup>Using data generated from the VCU calibration, we substituted the growth rate of the Solow residual in the SVAR in place of the growth rate of labor productivity. We found some reduction in the degree of upward bias in hours worked, though confidence bands around that variable widened somewhat. The slow adjustment of capacity utilization to both technology and non-technology shocks translates into slow adjustment of the Solow residual. Thus, the qualitative problems that contribute to bias that we have identified using labor productivity would appear to be relevant for this alternative approach.

estimated response of hours is at least 0.5 percentage point above the true response uniformly in the first year. In the VCU model, there is a 44% percent chance of such an occurrence. This result is interesting, since it highlights there is a high probability that a researcher might falsely conclude that hours worked rise markedly after a technology shock, when in fact they do not respond much. Similarly, the 75% probability of a large error in estimating the investment response reported in Table 4 mainly reflects the high probability of over-estimating investment.

The row of Table 7 labelled “Benchmark VCU” shows that the contribution of technology shocks to the variance of HP-filtered output is 28 percent for the VCU calibration. Unlike the benchmark FCU calibration, the SVAR’s median estimate of this contribution of 46 percent overestimates the importance of technology shocks. This overestimate reflects that the SVAR’s under-prediction for the response of labor productivity is more than offset by its over-prediction for the response of hours worked. Overall, the SVAR still yields an imprecise estimate of this contribution, but one that is more likely to be overstated than understated.

In the benchmark VCU calibration  $\rho_\chi$ , the parameter governing the persistence of the labor-supply shock, was set at the same value as our estimate for the persistence of the labor-tax shock. We also recalibrated the model with  $\rho_\chi = 0.5$ . Table 2 shows the selected moments for this calibrated version of the model, labelled “VCU with  $\rho_\chi = 0.5$ ”. Tables 3-6 report the relative distances and probability measures for this version of the model. As shown by the relative distance measure, the upward bias of the response of hours worked virtually disappears, and the probability of a large miss for labor productivity and output is slightly smaller. In addition, because the upward bias of hours worked is smaller in this case than in the benchmark VCU calibration, the SVAR tends to underestimate the importance of technology shocks to output volatility (see Table 7).

## 5 Incorporating Nominal Rigidities

In this section, we incorporate nominal wage and price rigidities into the FCU variant of the real business cycle model analyzed above. In particular, we assume that nominal wages and

prices are set in Calvo-style staggered contracts in a framework similar to that discussed in Erceg, Henderson, and Levin (2000), and utilized in a large related literature. Our comparison of this model to those considered above reinforces our finding that the bias in the estimated responses to a technology shock is model-dependent.

The inclusion of nominal rigidities into the model requires us to specify a monetary policy rule, and to calibrate several new parameters that are associated with the monopolistic competition-sticky price framework.<sup>27</sup> We assume that the central bank adjusts the short-term nominal interest rate in response to the four-quarter average inflation rate and to the current and lagged output gaps:

$$i_t = \gamma_i i_{t-1} + \gamma_\pi \pi_t^{(4)} + \gamma_{y,1} q_t + \gamma_{y,2} q_{t-1} + \epsilon_{it} \quad (22)$$

where the four-quarter average inflation rate  $\pi_t^{(4)}$  is defined as  $\pi_t^{(4)} = \frac{1}{4} \sum_{j=0}^3 \pi_{t-j}$ ;  $q_t$  is the output gap; and  $\epsilon_{it}$  is a monetary policy innovation (note that constant terms involving the inflation target and steady-state real interest rate are suppressed for simplicity). This form of the interest rate reaction function was estimated by Orphanides and Wieland (1998), who found that it provided a good in-sample fit over their 1980:1-1996:4 estimation period. Using the estimated parameter values found by Orphanides and Wieland (1998), we set  $\gamma_i = 0.795$ ,  $\gamma_\pi = 0.625$ ,  $\gamma_{y,1} = 1.17$ ,  $\gamma_{y,2} = -0.97$ , and the standard deviation of  $\epsilon_{it}$  is 0.0032.

The output gap  $q_t$  is the difference between the log of actual output,  $y_t$ , and the log of potential output,  $y_t^*$ . We assume that the central bank's potential output measure only responds to the level of technology and the capital stock so that

$$y_t^* = (1 - \alpha) z_t + \alpha k_t + (1 - \alpha) n, \quad (23)$$

where  $z_t = \log(Z_t)$ ,  $k_t = \log(K_t)$ , and  $n$  denotes the steady state value of  $n_t = \log(N_t)$ .<sup>28</sup>

---

<sup>27</sup>As is typical in this literature, we assume that money is separable in the utility function. Thus, with monetary policy specified by an interest rate rule, the equilibrium dynamics of our model can be determined independently of the quantity of money.

<sup>28</sup>An alternative measure of potential output commonly used is the level of output that would prevail if prices and wages were flexible. Relative to the results for our measure of potential output, this alternative measure would lead to results that are more similar to the RBC model already discussed.

Thus, the policy rule responds to the deviation of hours worked from its steady state level since  $q_t = (1 - \alpha)(n_t - n)$ .

The wage and price contracts have a mean duration of four quarters, and the wage and price markups both equal  $1/3$ . Given that all of the variation in the Solow residual reflects variation in technology, we set  $\mu_z = 0.0037$  and  $\sigma_z = 0.0148$ , as in the RBC model with FCU. Following the calibration methodology described earlier, we set the innovation variance of the labor supply shock,  $\sigma_\chi = 0.041$ , so that it allows the model to match exactly the observed ratio of the standard deviation of (HP-filtered) hours worked relative to output. The capital adjustment cost parameter,  $\phi_K = 5.05$ , and is set to match observed investment volatility. Finally, all other parameters are set identically to the values reported in Table 1.

The solid lines of Figure 6 show the true response of labor productivity, hours worked, consumption, and investment to a technology innovation in the model with nominal rigidities, while the dashed line shows the mean response from the SVAR. As in the benchmark FCU calibration, the mean responses of labor productivity, consumption, and investment understate the true response, and this downward bias in part reflects the slow adjustment of capital.

Another reason for the bias is that the estimated technology shock is contaminated by other shocks and in particular by the labor supply shock. This compositional bias is most apparent in the “near-term” response of investment. The true response of investment to a positive labor supply shock is initially negative in the model with nominal rigidities, because monetary policy is tightened in response to a such a shock. This tightening occurs because a positive labor supply shock puts upward pressure on hours worked and the output gap. Because the SVAR confounds positive labor supply shocks with technology improvements, the median estimated response of investment to a technology shock turns out to be seriously downward biased (see Table 3).

Thus, while labor-supply shocks contribute to pronounced upward bias in the estimated response of investment to a technology shock in the VCU model, they induce marked downward bias in the model with nominal rigidities. This comparison corroborates how bias in the estimated responses to a technology shock is model-dependent, and may be quite sensitive to



the responses of a variable to non-technology shocks.

The dotted lines in Figure 6 show the 90 percent confidence intervals of the impulse responses from the SVAR. It is clear that there is a wide band of uncertainty around the estimated responses of labor productivity, consumption, and investment. This is also confirmed in Tables 4-6, which report the probability that the estimated IRF is uniformly far from the true response for four, eight, and twelve quarters.

In Figure 6, there is a small initial decline in hours worked following a positive technology shock. There is a substantial likelihood that a researcher would fail to detect this initial decline based on estimates from the SVAR, since there is a 24% chance that the estimated initial response of hours is positive. However, in the model with nominal rigidities, the reliability of the SVAR's estimated response of hours is sensitive to the monetary policy rule. As pointed out in Galí, López-Salido, and Vallés (2003), if the monetary policy rule responds aggressively to stabilize output, hours worked can fall sharply in response to a positive technology shock. In this case, the structural VAR's estimated response of hours worked may be reliable enough to identify the initial fall in hours worked.

## 6 Evaluating confidence intervals

The confidence intervals reported in Figures 1 to 6 are constructed by generating 5000 data samples from the DGE model. In practice, researchers only observe one sample from the data-generating process, and therefore, it is interesting to assess the confidence bands produced by a researcher limited in this way. We estimated such confidence intervals following the non-parametric bootstrap method discussed in Runkle (1987). More specifically, for a given data sample generated from the DGE model, we bootstrapped 1000 new artificial samples using the point estimates of the VAR parameters as pseudo-true values. We applied our identification procedure to each of these new artificial samples, which allowed us to calculate point-wise confidence bands with a 90% probability content. This exercise was repeated for 1000 replications of data from the DGE model, and we tallied whether the bootstrap confidence bands covered

a variable’s true response in each period after the shock between 1 and 12. This exercise gave us 12 effective coverage rates for each variable in the VAR.

Table 9 reports the average coverage rates over the first 12 periods for labor productivity, output, hours worked, consumption, and investment. As shown in the row labelled “Benchmark FCU”, for labor productivity the bootstrap 90% confidence interval includes the true response only 35% of the time on average between lag 1 and 12. This poor coverage is partly attributable to the downward-bias of the estimated impulse responses. The coverage of these confidence intervals improves markedly for consumption and investment, as the bootstrap confidence intervals include the true response on average over 70% of the time. For hours worked, the coverage of the bootstrap confidence intervals exceeds 90%.

The effective coverage rates generally improve for the benchmark VCU calibration. However, this reflects that conventional confidence bands are generally extremely wide relative to the FCU model. Thus, while there is substantial bias in hours worked and investment, coverage rates are still quite high. By contrast, in the model with nominal rigidities, there is an overall deterioration in the coverage rates relative to the VCU version of the RBC model. The conventional confidence bands are narrower, and hence bias in the estimated responses tends to translate into poor effective coverage rates.

Kilian (1998) suggested a way to correct the bootstrap procedure for bias in the VAR estimates, which he called the “bootstrap-after-bootstrap” method. Before bootstrapping to determine the confidence intervals, this method performs an initial bootstrap to determine the bias.<sup>29</sup> As Table 9 shows, this correction procedure yields some improvement in the coverage of the confidence intervals.

In general, our results show that the effective coverage of confidence intervals for a given probability content is model- dependent. In models for which small-sample bias is large, hypothesis testing based on standard confidence bands may lead to more frequent rejections of

---

<sup>29</sup>For a given replication of the data from the DGE model, we used 1000 artificial datasets generated from the point estimates of the VAR parameters to determine their bias, the first step of Kilian’s procedure. Then 2000 new samples were bootstrapped, taking the bias-corrected point estimates as the pseudo-true values. As in Sims and Zha (1998), this bootstrap procedure was conditional on the initial observations.

a true null hypothesis than desired. As an alternative, a researcher could check whether the impulse response obtained from applying the Galí identifying scheme on actual data is included in the model-consistent confidence intervals (such as shown in Figure 1). This comparison is valid under the null hypothesis that the model at hand is the data generating process. The advantage of this approach is that it explicitly recognizes and adjusts for model-dependent small-sample bias.

## 7 Conclusion

While identifying technology shocks and their effects is a difficult task, our analysis suggests that Galí’s methodology is a useful tool. In particular, responses derived from Monte Carlo simulations of alternative models are qualitatively similar to the true responses. Quantitatively, however, we found considerable bias and spread in the estimated impulse responses. These deficiencies reflect small-sample problems associated with the long-run identifying scheme. We showed that the magnitude of the bias and spread depend on various characteristics of the true model, including the speed of capital adjustment, the persistence of non-technology shocks, and their size relative to technology shocks. We conjecture that a model that allowed for relatively smaller technology shocks or more endogenous persistence might pose more difficulty for the long-run identifying scheme than the models considered in this paper. However, such an investigation is a task for future research.<sup>30</sup>

Our analysis has two lessons for a researcher interested in comparing the responses to a technology shock derived from a theoretical model to those from applying Galí’s identification scheme to actual data. First, it is important to adjust for the small-sample bias rather than simply comparing the model’s impulse response to the estimates from the structural VAR. Second, given that the effective coverage rates of confidence intervals constructed by standard methods may be low, we suggest that researchers check whether the estimated response lies within the confidence bands generated by their model.

---

<sup>30</sup>Moreover, it would be interesting to compare the performance of the Galí identification scheme with that of the intrinsically model-dependent approach of Basu, Fernald, and Kimball (1998).

## References

- Altig, D., L. J. Christiano, M. Eichenbaum, and J. Lindé (2003). Technology Shocks and Aggregate Fluctuations. Mimeo, Northwestern University.
- Anderson, G. and G. Moore (1985). A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models. *Economic Letters* 17, 247–52.
- Basu, S., J. Fernald, and M. Kimball (1998). Are Technology Improvements Contractionary? Federal Reserve Board, International Finance Discussion Papers No. 625.
- Blanchard, O. J. and C. M. Kahn (1980). The Solution of Linear Difference Models under Rational Expectations. *Econometrica* 48(5), 1305–1312.
- Blanchard, O. J. and D. Quah (1989). The Dynamic Effects of Aggregated Demand and Supply Disturbances. *American Economic Review* 79(4), 655–673.
- Burnside, C. and M. Eichenbaum (1996). Factor Hoarding and the Propagation of Business Cycle Shocks. *The American Economic Review* 86(5), 1154–1174.
- Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2003). What Happens after a Technology Shock. NBER Working Paper No.w9819.
- Christiano, L. J. and J. D. Fisher (1995). Tobin’s  $q$  and Asset Returns: Implications for Business Cycle Analysis. NBER Working Paper No.5292.
- Cogley, T. and J. M. Nason (1995). Output Dynamics in Real-Business-Cycle Models. *American Economic Review* 85(3), 492–462.
- Cooley, T. and M. Dwyer (1998). Business Cycle Analysis without much Theory: A Look at Structural VARs. *Journal of Econometrics* 83, 57–88.
- Erceg, C. J., D. W. Henderson, and A. T. Levin (2000). Optimal Monetary Policy with Staggered Wage and Price Contracts. *Journal of Monetary Economics* 46, 281–313.
- Evans, C. L. (1992). Productivity Shocks and Real Business Cycles. *Journal of Monetary Economics* 29, 191–208.

- Faust, J. and E. M. Leeper (1997). When Do Long-Run Identifying Restrictions Give Reliable Results? *Journal of Business and Economic Statistics* 15(3), 345–353.
- Fisher, J. D. M. (2002). Technology Shocks Matter. Mimeo, Federal Reserve Bank of Chicago.
- Francis, N. and V. A. Ramey (2003). Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited. Mimeo, University of California at San Diego.
- Galí, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review* 89(1), 249–271.
- Galí, J., J. D. López-Salido, and J. Vallés (2003). Technology Shocks and Monetary Policy: Assessing the Fed’s Performance. *Journal of Monetary Economics* 50, 723–743.
- Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review* 78(3), 402–417.
- Hall, R. E. (1997). Macroeconomic Fluctuations and the Allocation of Time. *Journal of Labor Economics* 15(1), S223–S250.
- Hansen, G. (1985). Indivisible Labor and the Business Cycle. *Journal of Monetary Economics* 16, 309–327.
- Jones, J. B. (2002). Has Fiscal Policy Helped Stabilize the Postwar U.S. Economy? *Journal of Monetary Economics* 49, 709–746.
- Kilian, L. (1998). Small-Sample Confidence Intervals for Impulse Response Functions. *The Review of Economics and Statistics* 80(2), 218–230.
- Killingsworth, M. and J. Heckman (1986). Female Labor Supply: A Survey. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 1 of *Handbooks in Economics Series*, pp. 103–204. Elsevier Science B.V.
- Lippi, M. and L. Reichlin (1993). The Dynamic Effects of Aggregate Supply and Demand Disturbances: Comment. *American Economic Review* 83, 644–652.
- Lutkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

- Orphanides, A. and V. Wieland (1998). Price Stability and Monetary Policy Ineffectiveness when Nominal Interest Rates are Bounded at Zero. Finance and Economics Discussion Paper No. 98-35, Washington DC: Board of Governors of the Federal Reserve System.
- Parkin, M. (1988). A Method for Determining Whether Parameters in Aggregative Models are Structural. *Carnegie-Rochester Conference Series on Public Policy* 29, 215–252.
- Pencavel, J. (1986). Labor Supply of Men: A Survey. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 1 of *Handbooks in Economics Series*, pp. 3–102. Elsevier Science B.V.
- Pencavel, J. (2002). A Cohort Analysis of the Association Between Work and Wages Among Men. *The Journal of Human Resources* 37(2), 251–274.
- Rogerson, R. (1988). Indivisible Labor, Lotteries and Equilibrium. *Journal of Monetary Economics* 21, 3–16.
- Runkle, D. E. (1987). Vector Autoregression and Reality. *Journal of Business and Economic Statistics* 5, 437–442.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics* 6(2), 461–464.
- Shapiro, M. and M. Watson (1988). Sources of Business Cycle Fluctuations. *NBER Macroeconomics Annual*, 111–148.
- Sims, C. A. and T. Zha (1998). Error Bands for Impulse Responses. *Econometrica* 67, 1113–1156.

## 8 Appendix

This appendix is divided into three sections. In the first, we discuss how the log-linear solution of our RBC model can be written as a VARMA(4,5) for the benchmark FCU and VCU calibrations. We also describe how we verified that this VARMA process is invertible. In the second section, we show results for different fixed lag-lengths of the VAR. In the third section, we discuss our error decomposition.

### 8.1 Writing the RBC Model as a VARMA(4,5)

The first step involves log-linearizing and solving the RBC model around its nonstochastic steady state. This allows us to express the log-linear decision rule for the economy's scaled capital stock,  $\hat{k}_{t+1} = K_{t+1}/Z_t$ , as a function of lagged capital,  $\hat{k}_t$ , and the five exogenous shocks,  $\tilde{\mu}_{zt}, \tilde{\tau}_{Kt}, \tilde{\tau}_{Nt}, \tilde{g}_t, \tilde{\chi}_{0t}$  in the RBC model, (where the tilde denotes that the variable is expressed in log deviation from its steady state value). Also, for convenience, we have defined  $\mu_{zt} = \log(Z_t) - \log(Z_{t-1})$  and rewritten equation (7) more generally as

$$\mu_{zt} = (1 - \rho_z)\mu_z + \rho_z\mu_{zt-1} + \sigma_z\epsilon_{zt}, \quad (24)$$

even though  $\rho_z = 0$ .

In our benchmark FCU and VCU calibrations,  $\rho_{\tau_K} = \rho_{\chi_0}$  so that the labor supply shock is observationally equivalent to the labor tax shock up to a scaling factor. Therefore, we can combine these two shocks into a composite shock,

$$\tilde{\tau}_{xt} = \tilde{\tau}_{Nt} + f\tilde{\chi}_{0t}, \quad (25)$$

where  $f$  is a scaling factor determined from the linear approximation to the model's solution. Using these four shocks, we can define  $S_t = (\tilde{\mu}_{zt}, \tilde{\tau}_{Kt}, \tilde{\tau}_{xt}, \tilde{g}_t)'$ .

The log-linear decision rule for the scaled capital stock can then be expressed as:

$$\tilde{k}_{t+1} = a_{kk}\tilde{k}_t + b_{ks}S_t, \quad (26)$$

where  $a_{kk}$  is a scalar and  $b_{ks}$  is a 4x1 vector of coefficients. We can also write hours worked, the consumption-to-output ratio, and investment-to-output ratio as a function of  $\tilde{k}_t$  and  $S_t$ , while

the growth rate of labor productivity is a function of  $\tilde{k}_t$ ,  $\tilde{k}_{t-1}$ ,  $S_t$ , and  $S_{t-1}$ . Therefore, the model's dynamics for  $X_t$ , the vector containing the variables in our VAR, can be expressed as:

$$\tilde{X}_t = C_1 \tilde{k}_t + C_2 \tilde{k}_{t-1} + D_1 S_t + D_2 S_{t-1}, \quad (27)$$

where  $C_1$  and  $C_2$  are 4x1 vectors and  $D_1$  and  $D_2$  are 4x4 matrices.

Using the log-linear decision rule for  $k_{t+1}$  to substitute the scaled capital stock out of the linear decision rules for labor productivity growth, hours, and the ratios of consumption and investment to output, we can express the linear dynamics of  $X_t$  as:

$$X_t = a_{kk} X_{t-1} + (B_0 + B_1 L + B_2 L^2) S_t \quad (28)$$

$$S_t = \rho S_{t-1} + \sigma \epsilon_t$$

where  $B_0 = D_1$ ,  $B_1 = C_1 B_{ks}$ , and  $B_2 = C_2 B_{ks} - a_{kk} D_2$ ;  $\rho$  and  $\sigma$  are diagonal 4x4 matrices whose respective elements contain the AR(1) coefficients and standard deviations of the innovations.

Finally,  $\epsilon_t = (\epsilon_{zt}, \epsilon_{\tau_K, t}, \epsilon_{\tau_x, t}, \epsilon_{gt})'$ .

It is convenient to rewrite the first equation in (28) as:

$$(I - a_{kk} L) X_t = \sum_{j=1}^4 (B_{0,c(j)} + B_{1,c(j)} L + B_{2,c(j)} L^2) S_{jt}, \quad (29)$$

where  $B_{0,c(j)}$  denotes the  $j^{th}$  column of  $B_0$ , and  $S_{jt}$  is the  $j^{th}$  shock in  $S_t$ . Because  $\rho$  and  $\sigma$  are diagonal matrices, we denote the  $j^{th}$  element along the diagonal of these matrices as  $\rho_j$  and  $\sigma_j$ , respectively. Using these diagonal matrices, we can substitute out  $S_t$  from equation (29) to write

$$\begin{aligned} \sum_{i=2}^4 (1 - \rho_i L) (I - a_{kk} L) X_t = \\ \sum_{i=2, i \neq j}^4 \sum_{j=1}^4 (1 - \rho_i) (B_{0,c(j)} + B_{1,c(j)} L + B_{2,c(j)} L^2) \epsilon_{jt}, \end{aligned}$$

or

$$a(L) X_t = b(L) \epsilon_t, \quad (30)$$



with  $a(L) = \sum_{i=0}^4 a_i L^i$  and  $b(L) = \sum_{i=0}^5 b_i L^i$ . In the above,  $a_0 = I_4$  and  $a_i$  for  $i = 1, 2, 3, 4$  are 4x4 matrices that depend on  $a_{kk}$  and  $\rho_j$  for  $j = 2, 3, 4$ . Also,  $b_0 = B_0$  and  $b_i$  for  $i = 1, 2, 3, 4, 5$  are 4x4 matrices that depend on the elements of  $B_0$ ,  $B_1$ , and  $B_2$  and  $\rho_j$  for  $j = 2, 3, 4$ . Note that  $a(L)$  and  $b(L)$  do not depend on  $\rho_1$  since  $\rho_z = \rho_1 = 0$ .

Lippi and Reichlin (1993) make the point that researchers fitting a VAR to the data would not be able to recover the underlying shocks, if the data generating process had a non-fundamental representation. Therefore, for our benchmark calibrations, we checked that our model implied a fundamental representation by verifying numerically that the polynomial  $\det(b_0 + b_1 z + \dots + b_5 z^5)$  has all roots strictly outside the unit circle. This condition ensures that the VARMA process in equation (30) is invertible and is a fundamental representation for  $X_t$  (see page 222 and page 456 of Lutkepohl (1991)).

## 8.2 Estimation with fixed lag length

Table A shows the probability of a large error over the first four quarters for different lag lengths,  $p$ . There is some modest improvement in the fit of the SVAR, as indicated by our probability measure, for smaller values of  $p$ . Still, the probability of a large miss for labor productivity is above 30 percent, and there is over a 20 percent chance of concluding that hours worked falls when in truth it rises.

## 8.3 Error Decomposition

We can decompose the error in estimating the response to a technology shock into two sources:

$$\hat{d}_{l,i} - d_{l,i}^* = (d_{l,i} - d_{l,i}^*) + (\hat{d}_{l,i} - d_{l,i}), \quad (31)$$

where  $\hat{d}_{l,i}$  denotes the estimated impulse response for  $i^{th}$  variable, at lag  $l$  for a given draw of data. Also,  $d_{l,i}^*$  denotes the impulse response from the DGE model, and  $d_{l,i}$  is the population estimate of the SVAR's impulse response. We compute  $d_{l,i}$  by using the log-linear solution of the DGE model to find the population estimates of  $A_i$ ,  $i = 1, 2, \dots, p$ , and use those estimates along with equation (17) to determine  $A_0$ .

**Table A.** Varying the VAR Lag Structure for the Benchmark FCU Calibration: Probability that Estimated Response is Uniformly Far From True Response Over First Four Quarters<sup>a</sup>

Experiment	Labor Productivity	Output	Hours	Consumption	Investment
Lag Length = 1	0.31	0.18	0.23	0.17	0.24
Lag Length = 2	0.31	0.19	0.22	0.18	0.24
Lag Length = 3	0.32	0.21	0.22	0.19	0.24
Lag Length = 4	0.34	0.23	0.23	0.21	0.25
Lag Length = 5	0.36	0.25	0.23	0.22	0.26
Lag Length = 6	0.38	0.27	0.23	0.24	0.27
Lag Length = 7	0.39	0.29	0.23	0.26	0.29
Lag Length = 8	0.42	0.32	0.24	0.28	0.30
Lag Length = 9	0.45	0.34	0.25	0.30	0.32
Lag Length = 10	0.47	0.37	0.26	0.32	0.34
BIC	0.34	0.22	0.23	0.20	0.26

<sup>a</sup>For all variables except hours worked, the probability that the estimated response lies at least 33% above or below the true response for the first four quarters. For hours worked, the probability that the estimated response is uniformly negative in the first four quarters.

This first source of error ( $d_{l,i} - d_{l,i}^*$ ) arises because the VAR we estimate is an imperfect approximation of the VARMA process implied by our models. The second source ( $\hat{d}_{l,i} - d_{l,i}$ ) reflects small-sample bias. As discussed in Faust and Leeper (1997), small imprecision in estimating  $A(L)$  can result in large errors in  $R(1)$ , the long-run response of  $X_t$  to unidentified innovations. The error in estimating  $R(1)$  then affects all the parameter estimates of the SVAR through the long-run identifying scheme.

To separate out the small-sample error that occurs from imposing the long-run restriction from the error of estimating  $R_l$ , note that

$$\hat{d}_{l,i} = \hat{R}_{l,r(i)} \hat{\alpha}, \quad (32)$$

where  $\hat{\alpha}$  denotes the finite-sample estimate of the first column of  $A_0$ ,  $\hat{R}_l$  is the finite-sample estimate of  $R_l$ , and the subscript  $r(i)$  denotes the  $i^{th}$  row of this matrix. It is important to recognize that  $\hat{\alpha}$  is implicitly a function of  $\hat{R}(1)$  through equation (17). We follow Faust and Leeper (1997) and decompose the small sample error of estimating the impulse response of variable  $i$  at lag  $l$  as

$$\hat{d}_{l,i} - d_{l,i} = (\hat{R}_{l,r(i)} - R_{l,r(i)})\tilde{\alpha} + \tilde{R}_{l,r(i)}(\hat{\alpha} - \alpha). \quad (33)$$

The matrices,  $\tilde{\alpha} = \frac{1}{2}(\hat{\alpha} + \alpha)$  and  $\tilde{R}_{l,r(i)} = \frac{1}{2}(\hat{R}_{l,r(i)} + R_{l,r(i)})$  are defined to lie halfway between the finite-sample estimates and the population estimates of the SVAR. In equation (33), the small sample error,  $\hat{d}_{l,i} - d_{l,i}$ , has been decomposed into two parts: the first emphasizing the error in estimating the reduced-form moving average term,  $R_{l,r(i)}$ , and the second emphasizing the error in estimating  $R(1)$  through the  $\alpha$  term.

Table 1: Parameters Values Common Across Calibrated Versions of Model\*

$\beta = 1.03^{-0.25}$	$\rho_g = 0.95$
$N = 1/3$	$\sigma_g = 0.016$
$\chi = 3$	$\rho_{\tau_K} = 0.97$
$\delta = 0.02$	$\sigma_{\tau_K} = 0.008$
$\theta = 0.35$	$\rho_{\tau_N} = 0.98$
$\mu_z = 0.0037$	$\sigma_{\tau_N} = 0.0052$
$g/y = 0.20$	$\rho_\chi = 0.98$
$\tau_K = 0.38$	
$\tau_N = 0.22$	

\* $N$  and  $g/y$  denote the steady state values of labor and the ratio of government consumption to output.

Table 2: Selected Moments and Parameter Values of Calibrated Versions of Model<sup>a</sup>

Moment	U.S. Data <sup>b</sup>	FCU				VCU	
		$\sigma_\chi = 0$	Benchmark	High LSE <sup>c</sup>	Sticky Prices and Wages	Benchmark	$\rho_\chi = 0.5$
$\sigma_y$	2.17	1.40	1.59	1.90	1.54	1.82	1.67
$\sigma_h/\sigma_y$	0.80	0.34	0.80	0.80	0.80	0.80	0.80
$\sigma_c/\sigma_y$	0.47	0.65	0.54	0.59	0.56	0.62	0.63
$\sigma_i/\sigma_y$	2.91	1.96	2.69	2.31	2.91	2.91	2.91
$\sigma_{\Delta S}$	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Parameter Values							
$\phi_K$		0	0	0	5.05	0.46	2.86
$\sigma_z$		0.0148	0.0148	0.0148	0.0148	0.0076	0.0084
$\sigma_\chi$		0	0.021	0	0.041	0.022	0.023

<sup>a</sup>All moments except  $\sigma_{\Delta S}$  were computed by first transforming the data using the HP-filter (with  $\lambda = 1600$ ).  $\sigma_{\Delta S}$  refers to the standard deviation of the growth rate of the Solow residual.

<sup>b</sup> $\sigma_y$  and  $\sigma_h$  were computed using BLS data on nonfarm business sector output and hours from 1958-2002.

$\sigma_c/\sigma_y$  and  $\sigma_i/\sigma_y$  were taken from Christiano and Fisher (1995) who used DRI data from 1947-1995.

<sup>c</sup>Refers to model with  $\chi = 0$  and  $\sigma_\chi = 0$ .

Table 3: Percent Distance Between Mean Estimates and True Impulse Responses<sup>a</sup>

Experiment	Labor Productivity	Output	Hours	Consumption	Investment
Benchmark FCU	32	22	48	17	17
with $\phi_k = 100$	39	30	78	24	20
with $\sigma_z = 0.0049$ (1/3X)	44	8	360	10	43
with Lower Persistence <sup>b</sup>	16	14	14	14	13
with $\rho_{\tau_k} \approx 1$	31	23	31	20	17
with $\sigma_{\tau_k} \approx 0$	36	22	74	18	15
with hours differenced	17	29	110	25	24
High LSE	29	23	12	22	21
Benchmark VCU	31	12	550	2	67
with $\rho_\chi = 0.5$	28	27	26	21	23
Sticky Prices and Wages <sup>c</sup>	43	38	NA	31	45

<sup>a</sup>Absolute value of percent difference between mean estimated response and true response averaged over first twelve periods.

<sup>b</sup>Lower persistence refers to the case where AR(1) parameters of non-technology shocks set to half the benchmark values.

<sup>c</sup>With sticky prices and wages, the true response of hours is close to zero in the second period.

Table 4: Probability that Estimated Response is Uniformly Far From True Response Over First Four Quarters<sup>a</sup>

Experiment	Labor Productivity	Output	Hours	Consumption	Investment
Benchmark FCU	0.34	0.22	0.23	0.20	0.26
with $\phi_k = 100$	0.48	0.35	0.34	0.34	0.27
with $\sigma_z = 0.0049$ (1/3X)	0.56	0.52	0.21	0.65	0.72
with Lower Persistence <sup>b</sup>	0.03	0.02	0.01	0.02	0.04
with $\rho_{\tau_k} \approx 1$	0.34	0.26	0.25	0.25	0.33
with $\sigma_{\tau_k} \approx 0$	0.37	0.20	0.21	0.20	0.22
with hours differenced	0.12	0.30	0.26	0.25	0.29
High LSE	0.30	0.23	0.10	0.21	0.24
Benchmark VCU	0.35	0.42	0.13	0.18	0.75
VCU with $\rho_\chi = 0.5$	0.33	0.28	0.20	0.23	0.36
Sticky prices and wages	0.38	0.42	0.15	0.31	0.45

<sup>a</sup>For all variables except hours worked, the probability that the estimated response lies at least 33% above or below the true response for the first four quarters. For hours worked, the probability that the estimated response is uniformly negative in the first four quarters.

<sup>b</sup>Lower persistence refers to the case where AR(1) parameters of non-technology shocks set to half the benchmark values.

Table 5: Probability that Estimated Response is Uniformly Far From True Response Over First Eight Quarters<sup>a</sup>

Experiment	Labor Productivity	Output	Hours	Consumption	Investment
Benchmark FCU	0.30	0.19	0.20	0.17	0.21
with $\phi_k = 100$	0.44	0.31	0.28	0.28	0.21
with $\sigma_z = 0.0049$ (1/3X)	0.47	0.44	0.18	0.57	0.62
with Lower Persistence <sup>b</sup>	0.02	0.008	0.001	0.01	0.01
with $\rho_{\tau_k} \approx 1$	0.29	0.22	0.22	0.21	0.27
with $\sigma_{\tau_k} \approx 0$	0.34	0.17	0.18	0.17	0.17
with hours differenced	0.09	0.27	0.21	0.22	0.24
High LSE	0.25	0.19	0.08	0.17	0.20
Benchmark VCU	0.32	0.33	0.11	0.14	0.66
VCU with $\rho_\chi = 0.5$	0.29	0.23	0.13	0.19	0.25
Sticky prices and wages	0.33	0.32	0.10	0.26	0.28

<sup>a</sup>For all variables except hours worked, the probability that the estimated response lies at least 33% above or below the true response for the first eight quarters. For hours worked, the probability that the estimated response is uniformly negative in the first eight quarters.

<sup>b</sup>Lower persistence refers to the case where AR(1) parameters of non-technology shocks set to half the benchmark values.

Table 6: Probability that Estimated Response is Uniformly Far From True Response Over First Twelve Quarters<sup>a</sup>

Experiment	Labor Productivity	Output	Hours	Consumption	Investment
Benchmark FCU	0.29	0.18	0.18	0.16	0.18
with $\phi_k = 100$	0.42	0.29	0.26	0.26	0.19
with $\sigma_z = 0.0049$ (1/3X)	0.39	0.36	0.17	0.49	0.55
with lower Persistence <sup>b</sup>	0.03	0.01	0.002	0.02	0.01
with $\rho_{\tau_k} \approx 1$	0.25	0.19	0.19	0.18	0.24
with $\sigma_{\tau_k} \approx 0$	0.31	0.15	0.16	0.15	0.14
with hours differenced	0.07	0.26	0.18	0.20	0.24
High LSE	0.22	0.16	0.07	0.15	0.17
Benchmark VCU	0.31	0.27	0.10	0.14	0.58
VCU with $\rho_\chi = 0.5$	0.28	0.21	0.11	0.18	0.22
Sticky prices and wages	0.32	0.27	0.09	0.23	0.20

<sup>a</sup>For all variables except hours worked, the probability that the estimated response lies at least 33% above or below the true response for the first twelve quarters. For hours worked, the probability that the estimated response is uniformly negative in the first twelve quarters.

<sup>b</sup>Lower persistence refers to the case where AR(1) parameters of non-technology shocks set to half the benchmark values.

Table 7: Contribution of Technology Shocks to Output Volatility<sup>a,b</sup>

Experiment	$RC_z$	$\hat{RC}_z^m$	$P(\hat{RC}_z < 0.5RC_z)$
Benchmark FCU	0.74	0.52	0.30
High LSE	0.81	0.58	0.34
Benchmark VCU	0.28	0.46	0.14
VCU with $\rho_\chi = 0.5$	0.38	0.28	0.37
Sticky prices and wages	0.77	0.40	0.49

<sup>a</sup> $RC_z = \sigma_{y|z}^2 / \sigma_y^2$  where  $\sigma_y$  denotes the unconditional standard deviation of HP-Filtered output and  $\sigma_{y|z}$  denotes the HP-Filtered standard deviation of output conditional only on technology shocks.

<sup>b</sup> $\hat{RC}_z$  denotes estimated relative contribution from the SVAR and  $\hat{RC}_z^m$  denotes the median estimate.



Table 8: Varying the Sample Size for the Benchmark FCU Calibration: Probability that Estimated Response is Uniformly Far From True Response Over First Four Quarters<sup>a</sup>

Number of Quarters	Labor Productivity	Output	Hours	Consumption	Investment
120 (10 years less)	0.62	0.51	0.25	0.42	0.45
180 (benchmark length)	0.34	0.22	0.23	0.20	0.26
200 (5 years more)	0.32	0.20	0.23	0.20	0.25
220 (10 years more)	0.30	0.18	0.23	0.18	0.24
260 (20 years more)	0.26	0.15	0.22	0.16	0.22
400 (100 years)	0.19	0.11	0.22	0.14	0.20
1000 (250 years)	0.08	0.05	0.17	0.09	0.10

<sup>a</sup>For all variables except hours worked, the probability that the estimated response lies at least 33% above or below the true response for the first four quarters. For hours worked, the probability that the estimated response is uniformly negative in the first four quarters.

Table 9: Effective Coverage Rates of 90% Confidence Intervals<sup>a</sup>

Model	Labor Productivity	Output	Hours Worked	Consumption	Investment
Standard Bootstrap <sup>b</sup>					
Benchmark FCU	0.35	0.55	0.92	0.74	0.73
Benchmark VCU	0.61	0.92	0.81	0.88	0.92
Sticky prices and wages	0.34	0.36	0.98	0.55	0.64
Bias Correction: Bootstrap after Bootstrap <sup>c</sup>					
Benchmark FCU	0.43	0.63	0.97	0.76	0.80
Benchmark VCU	0.70	0.95	0.92	0.90	0.95
Sticky prices and wages	0.40	0.48	0.89	0.65	0.72

<sup>a</sup>Average of point-wise coverage rates from lag 1 to 12.

<sup>b</sup>See Runkle (1987).

<sup>c</sup>See Kilian (1998).

Figure 1: The Effects of Technology Shocks in the Benchmark FCU Calibration

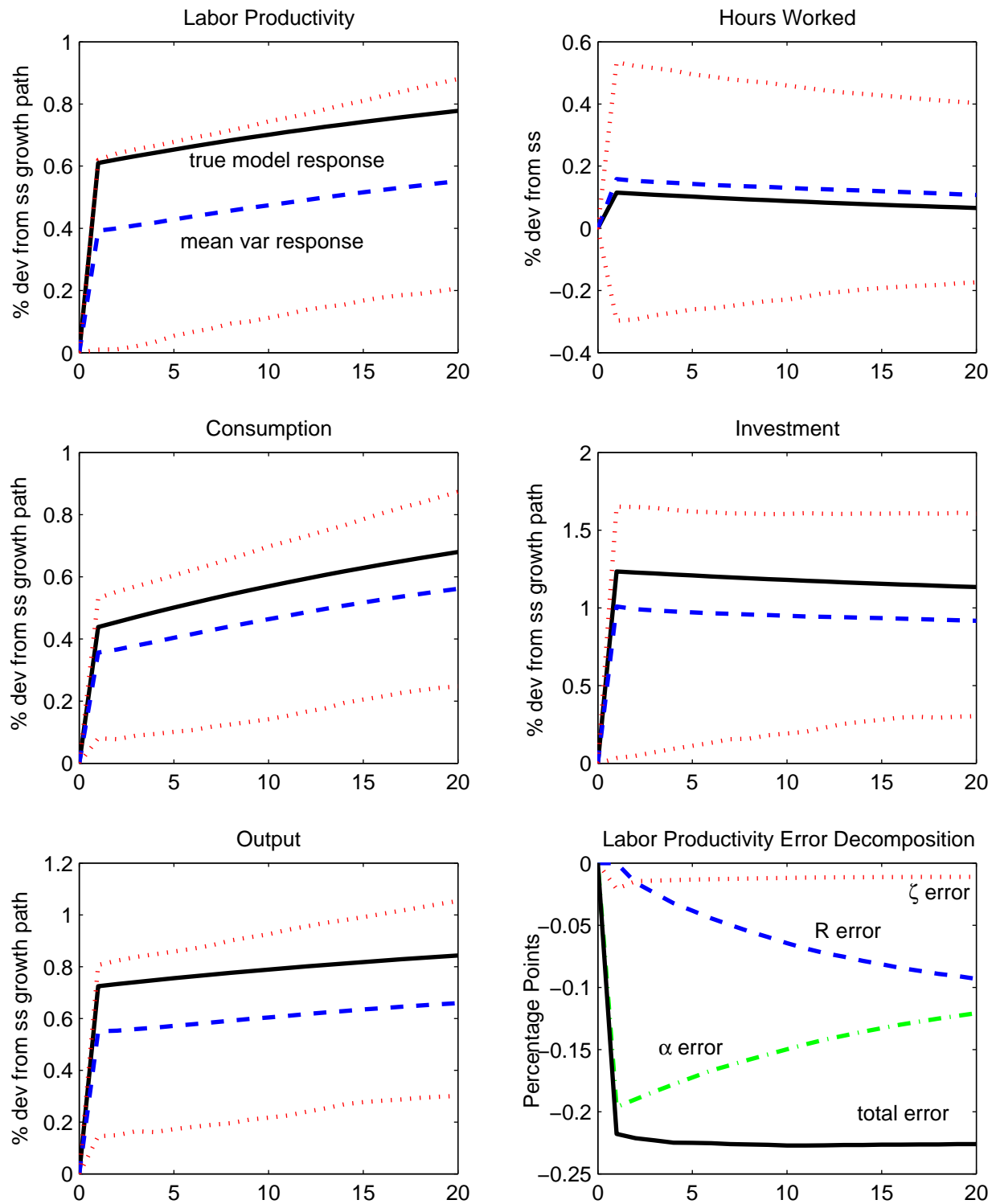


Figure 2: The Effects of Technology Shocks in the Benchmark FCU Calibration with  $\sigma_z = 0.0049$  (1/3X)

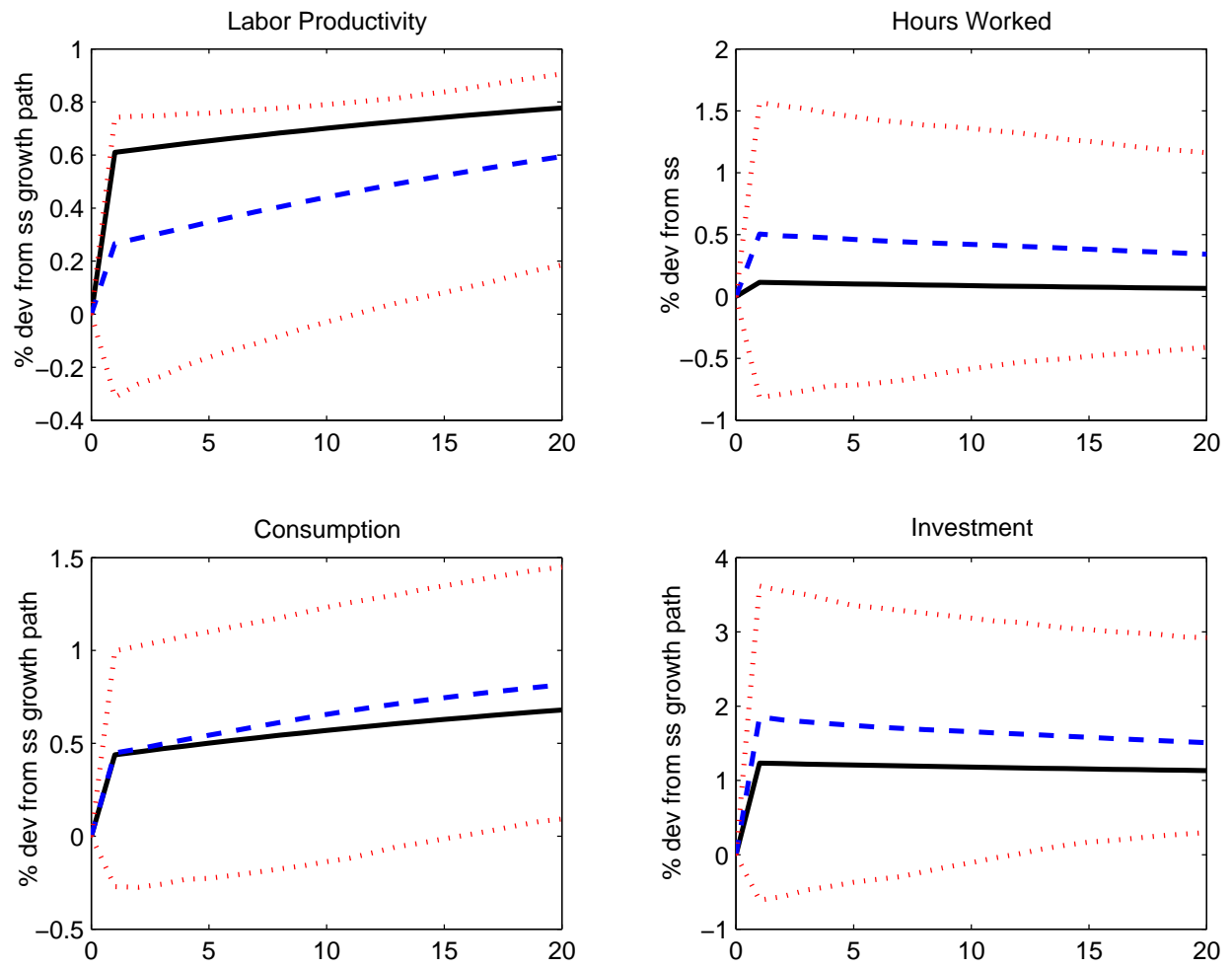


Figure 3: The Effects of One Standard Deviation Shocks in the Benchmark FCU Calibration with  $\sigma_z = 0.0049$  (1/3X)

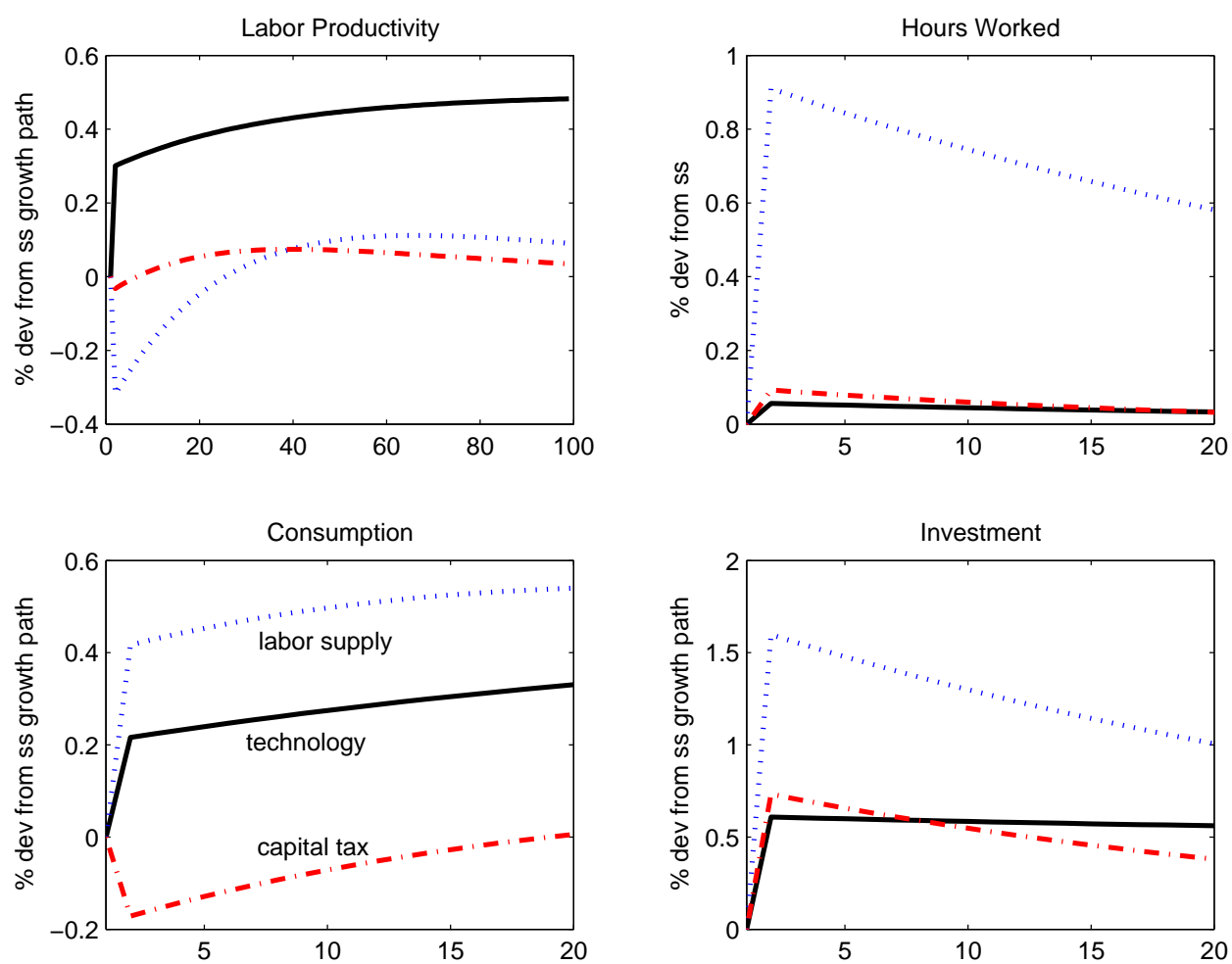


Figure 4: The Effects of Technology Shocks in High LSE Calibration

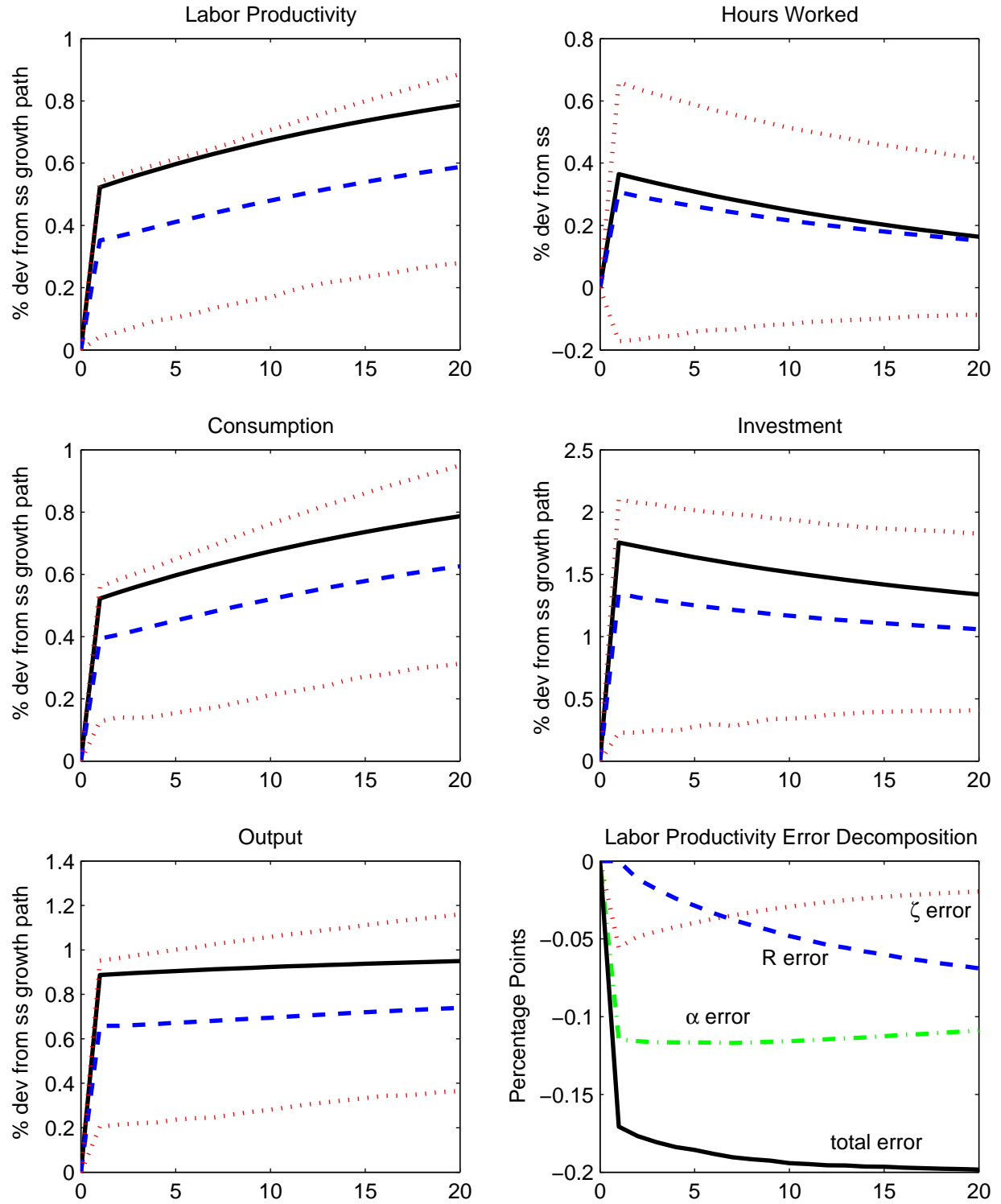


Figure 5: The Effects of Technology Shocks in the Benchmark VCU Calibration

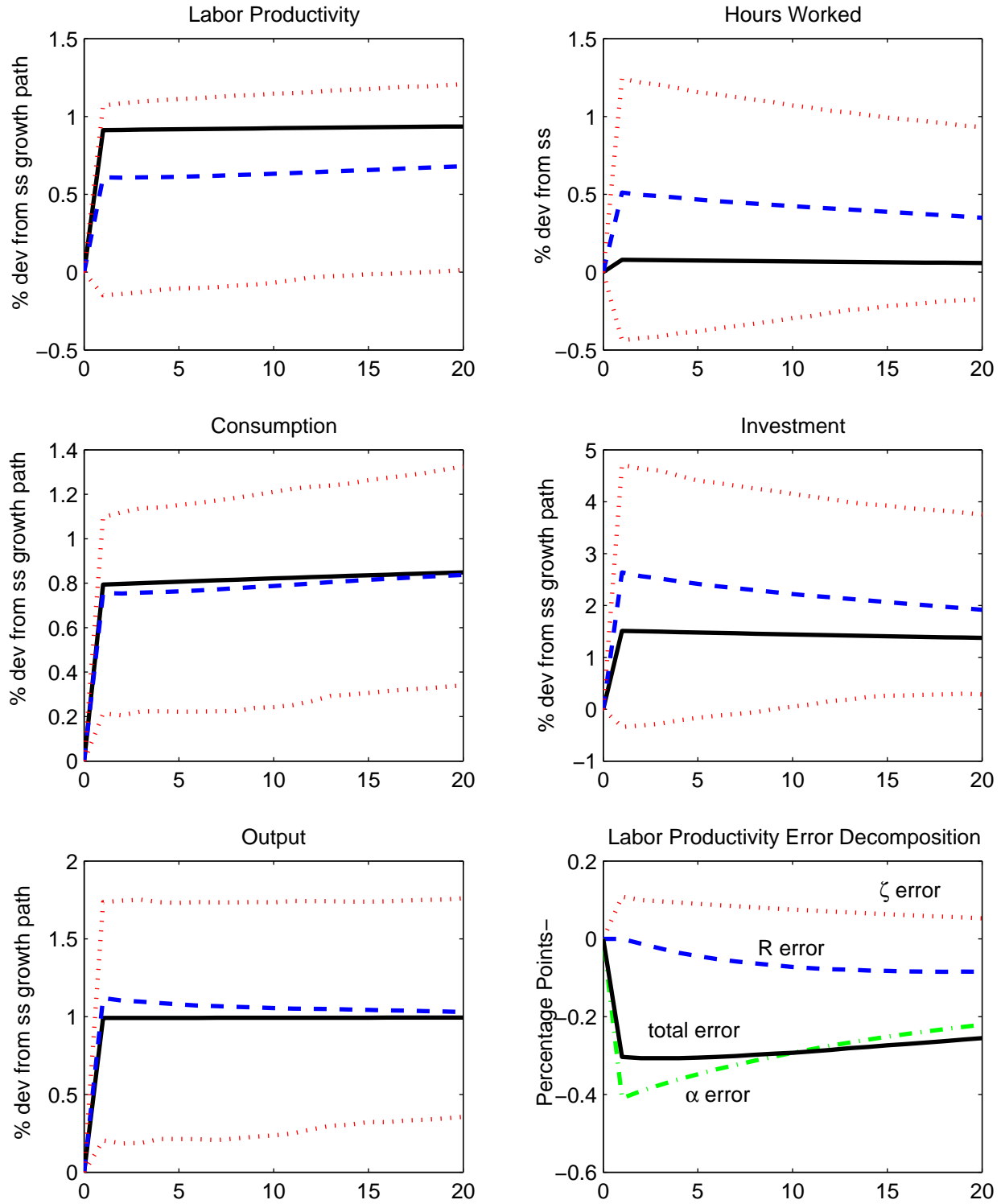


Figure 6: The Effects of Technology Shocks in Model with Sticky Prices and Wages

